

## Basic oral language documentation<sup>1</sup>

D. Will Reiman  
*SIL International*

Since at least 1992, when Michael Krauss presented the topic of language endangerment in *Language*, linguists have been wrestling with the problem that languages are disappearing from the earth faster than they can be satisfactorily documented. This paper advocates a methodology for documenting languages that minimizes the use of high-cost means of recording comments on recorded language data (written annotation), focusing instead on making low-cost means (oral annotation) more effective. I present here a brief history of the origins of the method, detail how the annotation process is executed, and evaluate its effectiveness in several dimensions. Finally, since it is an emerging technique, I will also discuss the directions in which the research on this methodology ought to develop.

### 1. INTRODUCTION.<sup>2</sup>

**1.1 BACKGROUND.** The urgency of language documentation requires that we do some serious rethinking of our priorities (cf. Krauss 1992:10). In the past, linguists have assigned the highest priority to the traditional products of descriptive linguistics namely, a grammar, a lexicon, and a corpus of interlinear texts. But we have been challenged by Himmelmann (1998) to recognize the distinction between language documentation (compiling and commenting on primary recordings of speech events) and language description (embodying the secondary results of analyzing the primary data and making generalizations). We have been further challenged by Woodbury (2003:45), who proposes that one could start the documentation process with purely oral techniques like producing “running UN style translations,” and, instead of transcribing everything, “starting with hard-to-hear tapes and asking elders to ‘respeak’ them to a second tape slowly so that anyone with training in hearing the language can make the transcription if they wish.”

**1.2 INDIGENOUS OPINIONS.** Documentation is not just an imposition by outside researchers. Many minority language communities want their language to be documented. During

---

<sup>1</sup> This article is an expansion of my presentation “Basic Oral Language Documentation” (Reiman, 2009).

<sup>2</sup> I have been blessed and honored to work on the Strategic Research Unit for Language Documentation, of SIL International. I have practiced and refined the ideas and theories of Gary Simons, with the strong technical advice of Brad Keating, an SIL ethnomusicologist. I must also especially acknowledge Mike Cahill, who provided comments and encouragement throughout this writing process. To them and the rest of the Unit, I must give much of the credit for what I write here. Of course, any errors and omissions are entirely my own.

the time I tested the Basic Oral Language Documentation (abbreviated as BOLD) methodology among the Kasanga,<sup>3</sup> various speakers said that their grown children leave the Kasanga area for jobs. When they return to visit, the grandchildren do not speak Kasanga and do not know Kasanga ways. They realize their numbers were larger, and now are small, and see that the language will most likely cease to be spoken soon. They also agree and appreciate it when told that their language and culture are valuable and worthwhile, and would like outsiders to learn and understand their language.

It is clear they see the value in documenting their language. If we carry the issue a step further by saying, “You should document your language!” we are likely to hear them comment, among other things, that “Our language is not written,” “We do not know how to write. Those of our people who write have moved to the city,” and “Learning to write/type/use a computer is expensive. It is time consuming. It has little benefit here in the forest.”

If the researcher and the speech community find value in a documentation project, and if both wish the ownership for the work and its products to rest in the hands of the speech community, the documentation project needs to be approached from the native-speaker point of view. The speech community can most benefit from a form of collaboration that meets mother-tongue (MT) speakers (preliterate language experts) on their own turf—a record of oral comments and annotations on the data of record. This paper will explain in detail these oral processes. I have found that oral processing enables MT speakers to compile and annotate their communities’ own language data corpora with minimal outside help.

**1.3 PRECURSORS OF BOLD.** At the 2003 annual meeting of the Linguistic Society of America, Tony Woodbury talked provocatively about NOT producing written interlinear glosses or written transcriptions for the ongoing Cup’ik documentation efforts.

We will ... produce running UN style translations ... We are also considering not transcribing everything—instead, ... asking elders to “respeak” [tapes] to a second tape slowly so that anyone with training in hearing the language can make the transcription if they wish (Woodbury 2003:45).

His notion of ‘respeaking’ caught the attention of my colleague Gary Simons. Gary immediately remembered a principal technique he practiced during his PhD research in the late ’70s.

[A] first tape recorder is used to play back the original text in short sections ... correspond[ing] to natural breaks in the text. After each section, the storyteller ... give[s] a translation of that section. [A] second ... recorder is left running ... to record both the original text and its translation... (Simons 1983:7).

---

<sup>3</sup> Spoken in and around Kanjandim, near São Domingos, Guinea-Bissau. ISO 639-3: ccj. Approx. 650 speakers.

Again, we have the idea of re-recording a text, but in this method, pausing the original at phrase breaks to insert translations orally.

Going back even further, we can find a similar technique described by Jacqueline Thomas (1976) in a field manual developed in West Africa for language research. Addressing the linguistic field worker, she gave them this advice to facilitate written transcription, and to make their job way easier:

Recording texts:

- a) Make an initial spontaneous recording of the narrator, ...
- b) Go over this spontaneous recording ... with a qualified speaker, in order to have it repeated sentence by sentence, in a careful, relatively slow, yet normal manner ...

What I am proposing is thus not entirely new, but it finds new space and purpose in the subfield of language documentation.

**1.4 MOTIVATIONS FOR BOLD.** Written forms of documentation often require painstaking study to get even a phonetic transcription, as well as an analysis for orthography. There are also costs in terms of transcription time, necessary personnel, finances, and accessibility in the case of pre-literate speech communities. For these reasons, my colleagues and I on SIL's Strategic Research Unit for Language Documentation have been dissatisfied with these written forms.

Many have used the oral annotation method with analog technology. This is a start; it can satisfy two of Himmelmann's (1998:11) three concerns—compiling language data and adding native speaker comments. This, however, does not adequately satisfy Himmelmann's third characteristic of good archiving, at least not in a timely and accessible way. A well archived corpus is able to yield specific data very quickly and easily. In nearly all cases, analog data records have failed on both accounts. As a consequence, we consider that an oral approach, with digital recording, can maintain or even improve the quality, sustainability, and shareability of documentation while decreasing time spent and training required, all with a smaller budget.

## 2. METHODOLOGY

**2.1 KINDS OF ORAL ANNOTATION.** There are three basic kinds of oral annotation:

**Careful speech.** Loosely referred to as “oral transcription,” it is meant to provide a clearer interpretation of the phones involved in the diction.

**Phrasal translation.** Each phrase in turn is translated into a language of wider communication.

**Analytical comments.** This is background information for mother tongue speech events, original to mother tongue speakers, that is made accessible to non-native speakers by a Language of Wider Communication. These are comments concerning things such as implied information, cultural knowledge, folk taxonomies, etc.

One can think of oral annotation as a novel way to encode the information linguists have been cataloguing for decades through interlinear record-keeping. A written documentation of a speech event will have multiple interlinear lines, as shown in Figure 1.

\ref	Argentina1 77					
\ori	papanakugumanani, inendirikuendaxikola					
\tx	papá	nakugumanani,	ine	ndiri	kuenda	xikola
\mr	papá	vamo-nos encontrar,	eu	estou	ir	escola
\trp	papá vamo nos encontrar eu vou para a escola.					
\tre	Father, I'll see you. I'm off to school.					
\nt	José remarks on the presence of the -ni suffix, ind. respect.					

\ref	Argentina1 78			
\ori	wamalakudya nkucewako			
\tx	wamala	kudya	nkhuce	wako?
\mr	acabaste	comer	merenda	sua?
\trp	acabaste de comer a sua merenda?			
\tre	You finished eating your <i>merenda</i> ?			
\nt	Estevão said ' <i>merenda</i> ' is unknown to the people of his region.			

FIGURE 1. Traditional text interlinearization.

Figure 1 demonstrates an interlinear text sample. Note especially the lines for original text phrases (\ori), transcriptions of those phrases (\tx), phrasal gloss lines also called “translation lines” (\trp and \tre), and lines for comments or notes (\nt). Each of these lines correlates to one of our three oral annotations:

\tx	‘transcription’	correlates to	careful speech
\trp, \trn	‘translation’	correlates to	phrasal translation
\nt	‘note’	correlates to	analytical comments

In this way, the oral documenter records the same basic elements that linguists have long been organizing through written means. The oral documenter, however, does not need to develop an orthography, train a keyboardist, nor set aside the hours required for written annotation work. They need only listen, press a pause button, and talk.

**2.2 OVERVIEW OF MY METHOD.** BOLD methodology combines aspects of each of the three scholars to whom I just referred.

By pausing the original speech event between phrases, the mother-tongue speaker can insert annotations into the audio stream. This is illustrated in Figure 2. We start with a recorder that we turn on and leave on (1). We then start playing the recording of our text (2), pausing at natural phrase breaks so the Mother-Tongue Annotator (3) has time to introduce comments on the just-previous phrase into the recording. Both signal sources are recorded, yielding a “turn taking” effect in the end product.

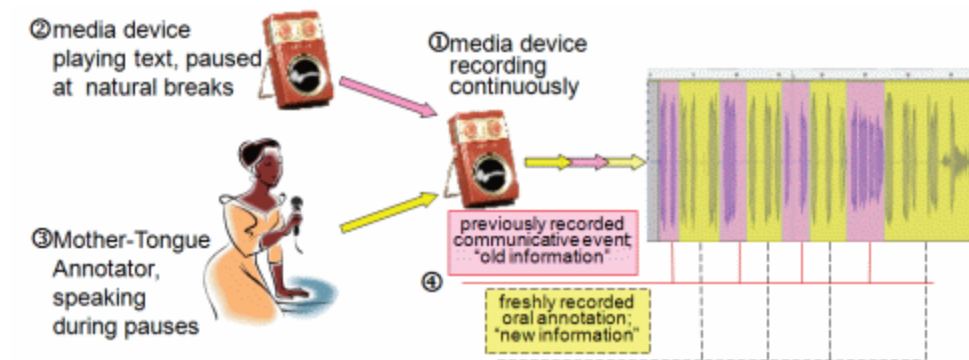


FIGURE 2. The Basic Oral Language Documentation process.

A researcher at any later time can then manually select the annotations that are required. However, if during equipment set-up one makes the minimal extra effort to send the original speech event's signal to one of the stereo channels—let us say the left channel—and the annotation stream to the other (in this case the right) channel, upon either aural or visual inspection, it is immediately obvious which part of the sound stream is the original speech event and which part is annotation, as seen in Figure 3.

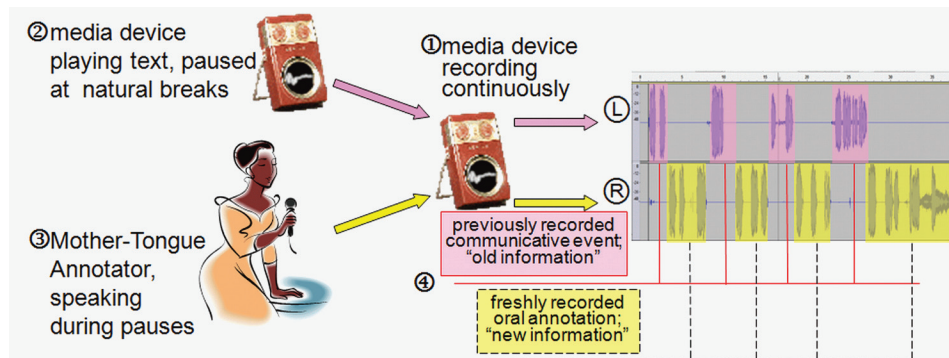


FIGURE 3. Channel-based segregation of old and new information.

With such channel segregation, distinctions between old and new information are now encoded digitally. This implies that this digital encoding is potentially exploitable by computer-based processing. Software could be developed, for example, that could export each original phrase and each annotation as separate WAV files, with each annotation labeled/named/tagged/linked to the appropriate originating file.

**2.3 AN EXAMPLE OF THE METHOD.** In October of 2007, I spent a week documenting a few aspects of the Kasanga language, whose speakers number about 650 (ISO 639-3 code: ccj). First, I recorded a communicative event. Most of my recordings were hortatory narratives. Figure 4 is the wave form from a 12-second sample from one of the texts I collected.

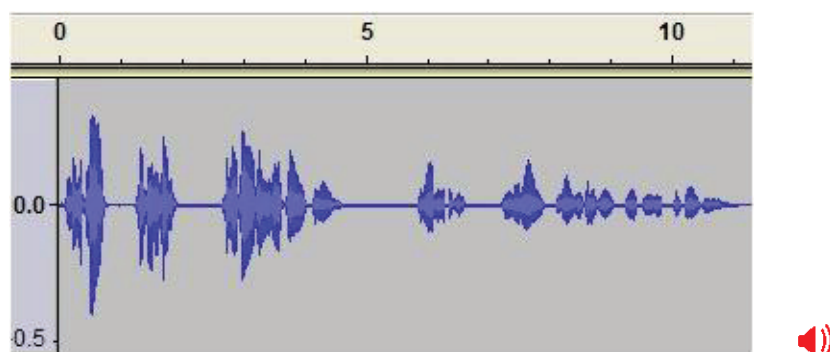


FIGURE 4. A 12 second sample of an original communicative event recording.

Next, I took that recording and played it into another stereo recorder. The Mother-Tongue Annotator listened to the process, pausing the playback on a phrase-by-phrase basis to repeat each phrase in careful speech. The re-recording of the original was kept on the left channel, and the careful repetition was kept on the right channel, to preserve and isolate the new information in its own channel. Here is this new creation:

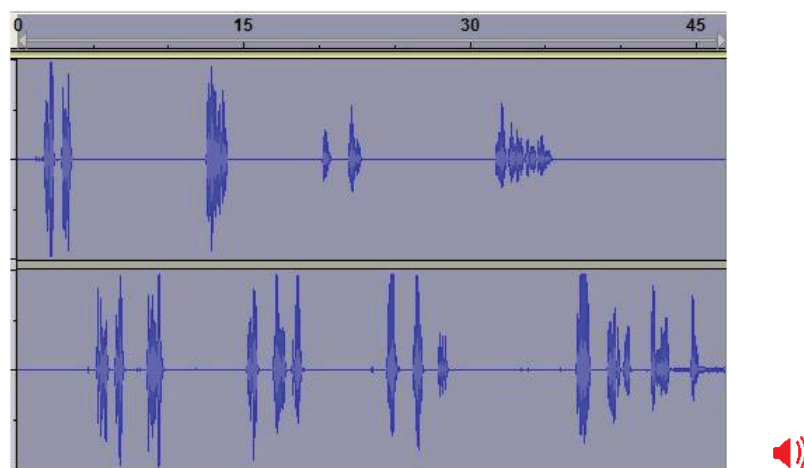


FIGURE 5. Careful speech annotations added to the original communicative event recording.

I used this process in multiple iterations, always isolating the old information on the left channel and isolating the new information on the right channel. Following is the waveform of the oral translation we made from the preceding recording:

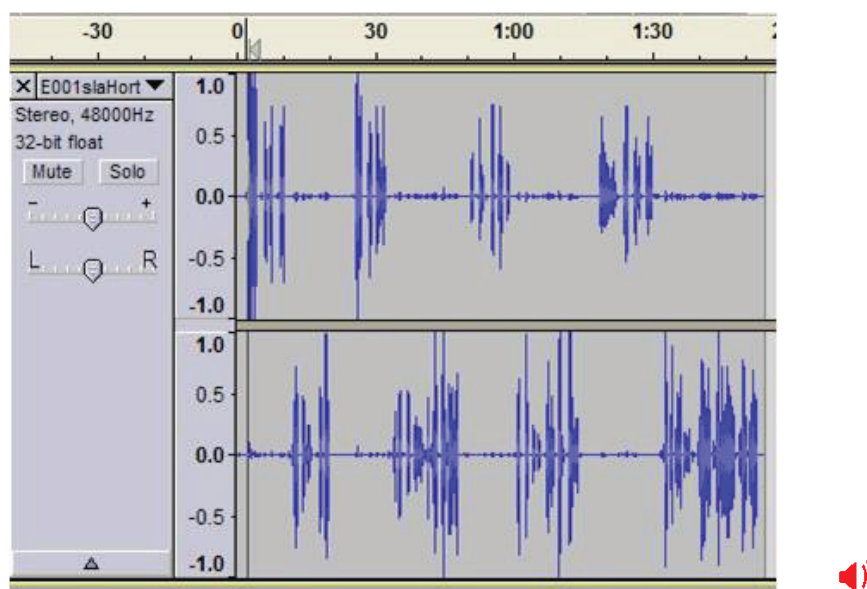


FIGURE 6. Phrasal translation annotations added to the preceding recording.

Note that the original recording and the “repetitions in careful speech” are combined onto the left channel as “old information” in Figure 6. The “old information” came from the stereo WAV file of the careful-speech annotation recording session, which I mixed to mono through a cable/adaptor setup. There are other ways to record the phrasal gloss, but this has a couple of advantages. First, this type of stereo data file can be recorded “in the field” with no computer support. Second, by having the annotator use the recording from the annotation session, the phrase break decisions are already made, and in the end, the phrases of careful speech and the phrasal glosses will refer to the same stretch of original text. If one were to use only the original recording, the gloss annotator would have to make fresh phrase-break decisions that might or might not line up with the decisions made by the careful-speech annotator.

The most critical skill to be acquired by the trainee was learning to appropriately use the pause button. Synchronizing this with a fast-moving recording is not always easy, but those I trained ended up handling it very well. I realized it was imperative to devote at least a couple of hours just to familiarize them with the technique.

This documentary workflow, then, could be represented as in Figure 7:



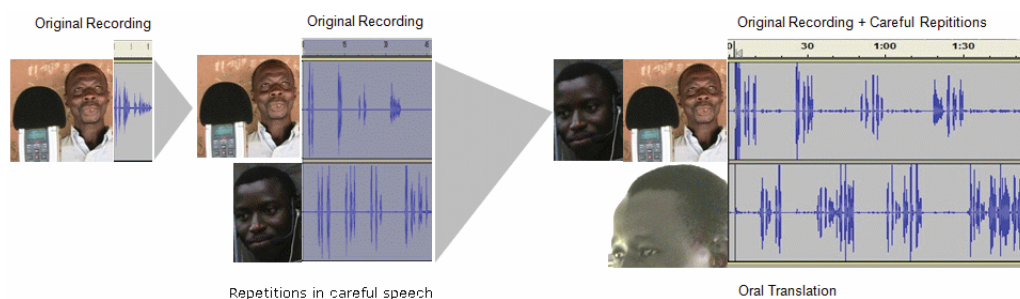


FIGURE 7. Documentary workflow

The wave forms shown follow my convention of old information being recorded to the left (upper) channel, and the new information being recorded to the right (lower) channel.

**3. WHO SHOULD DO WHAT, WHERE SHOULD THEY DO IT, AND HOW?** In my trial project with the Kasanga, it became apparent in just a few days that collecting original communicative events, and collecting annotations on those recordings, required very different recording environments, personnel, and equipment.

**3.1 LOCATIONS FOR EACH STEP.** The original recordings had to be collected in their natural and authentic environment, even at the expense of recording quality. These recordings should be made in the heartland of the speech community (along with the chickens, goats, dogs, and crying children) near the cultural insiders, which often means being away from studio conditions and grid power. Oral annotations—both careful repetitions and translations—do not require an authentic environment. These benefit most from a quiet place that allows the annotator to listen closely without distraction and gives the recordist an opportunity to collect a very high quality, noise-free data stream. Grid power will also greatly facilitate annotation and data management workflows.

**3.2 PERSONS FOR EACH STEP.** Just as different recordings are best made in different locations, it is also essential to have persons with specific skill sets. Original recordings demand the most expert performers, which often means older people who have lived long periods in the rural heart of their ethnic community. Such culturally experienced persons are preferred, in spite of weak voices or missing teeth. Oral annotations of careful speech should avoid persons with poor enunciation, if at all possible, choosing instead people who hear and speak the mother tongue extremely well and are not intimidated by the digital tools they will be handling. Oral translations should target yet another set of talents. An oral translator does not necessarily have to be the best speaker of the language of study, but rather must be an excellent speaker of the language of wider communication, while still possessing a very good ear for the language of study. This criterion is often best met by younger community members, who may not possess high-level speaking ability in the language of study, nor a commanding knowledge of their people's culture and cultural heritage. Moreover, in the areas where I have worked, these people typically live nearer the towns and cities.



**3.3 EQUIPMENT FOR EACH STEP.** High quality recording equipment, with its bulk, wires, and need to be close to the performers, is usually too intrusive and culturally inappropriate for recording original communicative events in situ. These recordings benefit from low-profile equipment, and low-profile techniques in using it, even at the expense of quality. Oral annotations, however, to be effective linguistic resources, should be recorded with quality equipment. Size, visibility, and power specifications must take a back seat.

**3.4 COORDINATING THE PEOPLE INVOLVED.** A well managed collaborative effort in language documentation can use many different combinations of linguists, field research technicians, mother-tongue speakers, and support workers to accomplish mutually agreed upon goals. There are not enough linguists to do all the tasks themselves, and it is costly and time consuming to develop the necessary number of fully trained linguists. If the linguistic community is to make headway in the rapid gathering of documentary corpora, field research technicians—persons partially trained in linguistics and field work—must be integrated into the work. Those in undergraduate programs in linguistics could easily “get their feet wet” in very practical ways by helping a seasoned field researcher gather language data. As I have been arguing, through the application of digital technology to oral annotation techniques, speech communities themselves can assist in the tasks of language data processing and, with brief training as short as half a day, will be able to gather the original data of communicative events with handheld recorders. Furthermore, people both within and outside the linguistic and/or speech communities could help with data management tasks and the initial analysis of some of the collected data.

Each of these groups has the potential to greatly increase the effectiveness of a documentation effort. I would like now to consider the increased effectiveness each group might bring to the task, both quantitatively and qualitatively.

Let us begin with that last suggestion of data bookkeepers. One need only be involved in a documentation project for a short time to see that data management becomes a huge prerequisite for guaranteeing the success of the project. Extra help here in the form of two or three computer keyboardists can go a long way toward keeping all the data organized and manageable. One must be careful, however, because the “garbage in, garbage out” principle is especially hard at work here. If those generating the recordings—whether a linguist, a technician, or a mother-tongue keyboardist—do not keep meticulous logs and records as they record, the bookkeeper will in short time not know what to do with the dozens of files automatically labeled by the recorder as ‘STE-001’ and the dozens labeled ‘MONO-002’, ‘MONO-003’, ‘MONO-004’, and so on.

The first line of action for preserving languages is to record the communicative events and lists of value. Certainly, this is the sine qua non for language documentation. Without recordings, there is no possibility of a language data corpus. Therefore, a research project planner might want to get sufficient equipment and persons to multiply the number of original communicative events being gathered per available field hour. A dozen mother-tongue speakers can be trained and sent out to village events with cheap dictation voice recorders (easily had for less than US \$70 each). Yet, while this could improve the content quality of the documentary corpus—you now have 12 versions of a specific folktale to choose from, instead of, say 2—this will not increase the speed with which the corpus can be collected. However, this is not where the most time, effort, and people are needed. Even with

a streamlined application of the BOLD methodology, eleven hours of annotation and file management are required for every hour of original data recording. To put it another way, for one half-day of recordings (3 hours—60 communicative events @ 3 minutes each), an annotator has to work nearly an entire week (at least 33 hours) to complete the annotation process. Hence the bottleneck, in terms of time, does not reside with recording original events, but with recording the necessary annotations.

Additionally, all the dictation recorders I have researched record using a lossy compression, most often WMA and MP3. I have never seen anything in that price range (< US \$70) that records at the level of the best-practice recommendation of 24-bit WAV. This therefore reduces the recording quality of the data collected. When push comes to shove, rather than spending \$700 on a dozen dictation recorders, I recommend spending the same money on four digital audio recorders, even the entry-level Zoom H2 at around \$140 each. This would equip two separate quiet rooms, rather than just one, thereby doubling annotation productivity.

**4. EVALUATION.** There are several reasons I like this method and there are various ways in which I see it held back by current limitations. Let me clearly state these pros and cons, and then go on to the steps that could alleviate the shortcomings.

#### **The Pros:**

##### **Rapid gathering of original communicative events**

It took approximately 11 hours to orally annotate each hour of original recordings. This is fast when one considers that the most favorable estimates of the number of work-hours sufficient to document with traditional keyboarded annotation begin at 20 hours of keyboarding for every hour of original recordings (Newman 2009). Most will calculate the time to be much longer.

##### **Easier engagement of the speech community**

Removing keyboard operations from the mother-tongue annotation process and reducing hardware training to a single-button control enables more preliterate speech community members to engage in the documentation process, especially at a much lower training cost.

##### **Relatively rapid production of intelligible speech segments, useful for analysis and for language learning**

When done well, the BOLD methodology yields intelligible phrases based on natural speech. This intelligibility is in two dimensions: the well articulated phrases give recognizable phones, and the oral translations of each phrase make the meaning understandable as well. These in combination could provide hundreds, if not thousands, of example phrases to be used by non-native speaker who want to learn the language.

##### **Aurally and visually discernable data groupings, with minimal extra cost**

By annotating the sound files orally with this “speak-to-the-other-channel-while-paused” method, the data is “automatically” organized in two ways. First, all the annotation is on a single track and can easily be extracted from

the “old information” on simple audio-file editors like Audacity.<sup>4</sup> Second, for any given annotation, one knows it is a comment on the immediately preceding segment in the other channel (see Figure 8). This can yield very quick visual searches with your favorite audio file editor, and it even aids searches by ear when one listens to the number of times the signal switches from the left channel to the right channel.

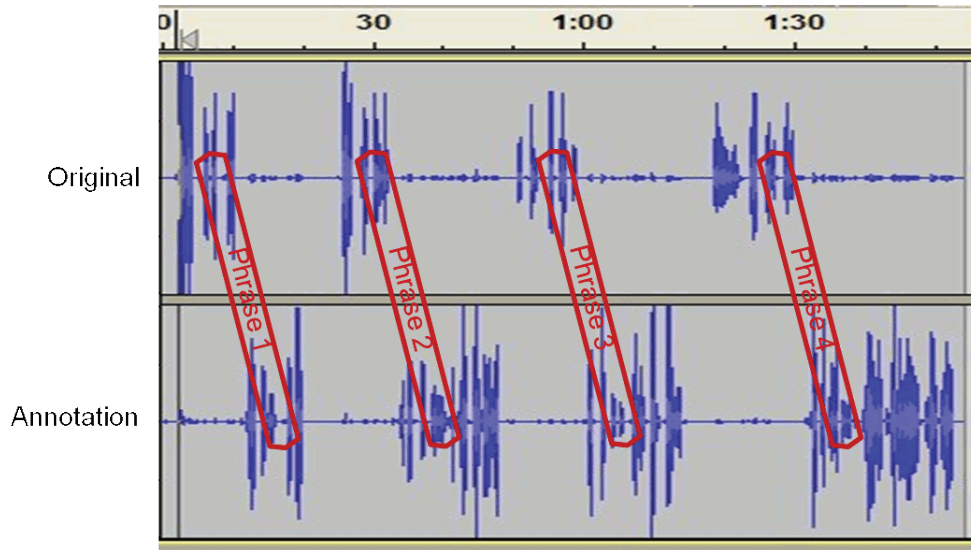


FIGURE 8. How annotations correlate to ‘old information’

#### The Cons:

**Without any editing, approximately 70% of the WAV files recorded contains non-meaningful “silence” or redundant information.**

One of the unfortunate problems with the method is that, since one channel is silent while one records on the other, 50% of the digital data being encoded to the WAV file on each channel is silence. Further, since one is making a real-time re-recording of the original file through an analog patch cable, duplicate information is produced on the new recording that is inferior in quality to that on the original recording. When totaling these re-recordings on a sample of my data, I found it to take up at least 20% of the data-space, bringing the total “junk signal” to more than 70%. When dealing with hundreds of 24-bit WAV files, that results in quite a bit of wasted digital storage space.

<sup>4</sup> Audacity can be downloaded at <http://audacity.sourceforge.net/>

### **No word-for-word translation, as recommended by Himmelmann (2009)**

Himmelmann (2009) suggests that a word-for-word translation is a vital element in a complete documentary corpus for a language. At this time, the BOLD methodology only yields glosses down to the phrase level. Our hope would be that this level, in combination with careful speech, would provide sufficient clues (hundreds in number) so as to act as a 'Rosetta Stone' for developing word-level glossing. Additionally, the elicitation of a minimum 1,000-item wordlist, with example sentences designed to give breadth and depth of syntactic information, should largely compensate for the lack of word-level interlinearized glossing.

A word-for-word oral gloss annotation is not out of the question. However, phrase-by-phrase oral translation is very common at public gatherings in the locations where I have conducted my research; word-for-word translation is not. Hence word-for-word oral glossing adds a layer of complexity to the training of annotators. However, one of my students integrated a word-for-word method in his field work with an annotator who had previous experience with entering more traditional, keyboarded word-level gloss lines. The method, level of success, efficiency, and breadth of applicability are currently being reviewed.

### **Many ways to record poorly**

Although a well-trained Mother-Tongue Annotator can produce very good documentation, a significant part of the job is art, rather than science. There are many ways to record annotations poorly. Syntactic phrase breaks do not always line up with phonological phrase breaks, increasing the difficulty of pausing the original recording at opportune times. Some digital audio recorders do not have an easy-to-use rewind control, requiring frequent intervention by the researcher. Practice helps to alleviate all of these problems. The Kasanga annotator I trained asked for two hours to simply "tool-up." Other annotators may need more time. Each new text should be listened to several times, such as when practicing the pauses, before a recording is attempted. This goes against the industrialist bent for efficiency, but, in the end it makes for a smoother, and arguably faster, recording process.

**5. NEXT EFFORTS.** All these problems stimulate the search for better methods, while all the advantages motivate persistence. Automated, computer-based assistance is very attractive and promises to be very helpful. However, it will benefit us to note a few challenges associated with computer-based (or computer-assisted) oral language documentation.

### **Challenge #1: Battery life**

Language documentation necessarily takes the researcher, literally, off the beaten path. Computer battery power that currently lasts at most eight to ten hours will not suffice when the researcher must travel off-the-grid for two weeks to two months. Technology predictions are optimistic for netbook power that can last a week (some figures are as optimistic as a month!). I believe the leap will

come. Until then, one must consider alternatives: drive a car and recharge from its battery? Carry solar panels? These will not work in all situations, but they will work in some. For now we can gather data with digital audio recorders, use a laptop solely for storing the sound files, and save the computer processing for when we get back to grid-power.

### **Challenge #2: Intrusiveness of technology**

It may be different for research in other parts of the planet, but I have come to agree with the opinion of my colleague Margaret Shrum: “People who are monolingual run from researchers and hide. They push the semi-bilingual up front to answer the questions. The truly bilingual get that way in schools outside their home areas and don’t come back unless they are hiding ...” (pers. comm.). In this way, the best speakers of a minority language are often hard to find, they avoid outside contact, and they are wary of unknown artifacts from other cultures. If I am fortunate enough to find such an expert and pull out my laptop during our first session together, the expert’s responses will in all likelihood become more reserved, and s/he may clam up altogether. If s/he is expected to interact through a keyboard, s/he may have an even stronger negative reaction.

Even so, I am desirous of computer assistance in oral language documentation. I see three broad areas of application as listed below.

**5.1 FILE MANAGEMENT.** No matter how one documents a language, if the method relies on digital resources, there will be hundreds of files in a corpus set, which will probably have a majority with names automatically generated by the recording device (STE-000, -001, -002 ... -398, -399, etc.) As I said of bookkeepers before, this tool would be an automated bookkeeper’s aid, as some sort of database management system. It would also keep track of the metadata for each file, including how different files relate to one another.

**5.2 SIGNAL PROCESSING.** Backing up a step from file management, if we generate files as described above that contain perhaps only 30% of meaningful data, an automated way to get rid of the non-meaningful segments (or, equivalently, keeping only the meaningful data) will be critical if this method is to be useful in cataloguing any significant number of world languages. Recorded at best-practice standards, an hour of recording consumes approximately a gigabyte in two-channel (stereo) mode, and stereo is how nearly all the WAV-recording digital audio recorders create their files, whether recorded with one microphone or two.

**5.3 DATA GATHERING.** Recordings could also be made directly to a computer. I want to talk about this with great caution, not only because of the intrusiveness of technology as I mentioned above, but also because there is a recording quality concern. Even though all laptops currently manufactured have a microphone jack, most are not good recording devices. The hard disc makes extra noise, and everything inside the laptop creates electromagnetic signals that an internal sound card without shielding will pick up. Most laptops do not have this shielding. Each laptop varies in its level of internal electromagnetic

noise, anywhere from hardly noticeable (count yourself lucky) to very distracting. For this reason, best-practice experts recommend using an external analog-to-digital interface if one intends to record directly to a computer. Portable external audio interfaces will cost upwards of US \$150. USB microphones of acceptable standards, to my knowledge, are not available for less than US \$500 at the time of this writing.

**5.4 ALL THAT BEING SAID...** In spite of these caveats about computer use, on my last text-gathering expedition—among the Manyawa (ISO 639-3: mny) during June and July of 2010 (data-set publication in-process)—I did indeed speed up the annotation process through the use of computers. Since our location for annotation had very reliable grid power, our team’s linguistic intern, Timothy Krug, and I were able to use currently available software to parse each text at phonological phrase breaks. This is not a straightforward process, but once we figured out a system for exporting the phrases through batch processes and ensured that our digital audio recorder would be able to play them (many brands, including our Zoom devices, are picky), things proceeded at a good pace. With each phrase being in a separate WAV file, and the digital audio recorder set up to stop playback after each file, the annotator no longer needed to make the phrase-break decisions or to be quick on the pause button. When the recording stopped, one would speak his/her annotation, click on the “next” key, and click “play.” The careful speech annotation and the phrasal gloss always referred to the same stretch of original text. Moreover, the phrasal gloss process went at least twice as fast as with the former methodology, since the annotator did not have to wait for both the original phrase and its careful repetition.

**6. CONCLUSIONS.** Since the vast majority of languages that need documentation have neither an orthography nor a literate community, the method I have described focuses on Oral Commenting, or *Oral Data Annotation*—: recording language information about a media file through audio (and/or video) recordings of MT speakers, which become tagged to the original media files. It is not without its drawbacks, but this method has as its advantage the capacity to engage the speech community quickly and early in the language documentation process. It can reduce the time required for documenting by attacking the real workload bottleneck of creating digital annotations (or comments, or metadata) from a preliterate speech community with no orthography. And since it is digital, it also has the capacity to make this data accessible throughout the digital world.

## REFERENCES

- Himmelmann, Nikolaus. 1998. Documentary and descriptive linguistics. *Linguistics* 36.: 165–191.
- Himmelmann, Nikolaus. 2009. Linguistic data types and documentary linguistics. Plenary address at the 1st International Conference on Language Documentation and Conservation [ICLDC-1]. University of Hawai‘i at Mānoa, Honolulu, Hawai‘i, 12–14 March, 2009. <http://hdl.handle.net/10125/5162>. (27 August, 2010.)
- Krauss, Michael. 1992. The world’s languages in crisis. *Language* 68(1). 4–10.
- Newman, Paul. 2009. Linguistic fieldwork as a scientific enterprise. Plenary address at the 1st International Conference on Language Documentation and Conservation [ICLDC-1]. University of Hawai‘i at Mānoa, Honolulu, Hawai‘i, 12–14 March, 2009. <http://hdl.handle.net/10125/5163>. (31 August, 2010.)
- Reiman, D. Will. 2009. Basic Oral Language Documentation. Paper presented at the 1st International Conference on Language Documentation and Conservation [ICLDC-1]. University of Hawai‘i at Mānoa, Honolulu, Hawai‘i, 12–14 March, 2009. <http://hdl.handle.net/10125/5071>. (31 August, 2010.)
- Simons, Gary F. 1983. Language variation and limits to communication. Dallas, Texas: Summer Institute of Linguistics.
- Thomas, Jacqueline M. C. 1976. *Enquête et description des langues à tradition orale. I. L’enquête de terrain et l’analyse grammaticale. II. Approche linguistique (Questionnaires grammaticaux et phrases) III. Approche thématique (Questionnaire technique et guides thématiques)*. Paris, SELAF, 2nd edn.
- Woodbury, Anthony C. 2003. Defining language documentation. In Peter K. Austin (ed.), *Language Documentation and Description* 1., 35–51. London: SOAS

D. Will Reiman  
will\_reiman@sil.org